Estimation du paramètre de dérive d'une diffusion sous des conditions d'irrégularité de la dérive.

Sandie Souchet

Rubrique : Statistique/Statistics

Titre : Estimation du paramètre de dérive d'une diffusion sous des conditions d'irrégularité de la dérive.

Title : Estimation of drift parameter under irregularity condition for drift.

Titre courant : Estimation paramétrique sous des conditions non régulières.

Sandie SOUCHET SAMOS, Université Paris 1, 90, rue de Tolbiac, 75013 PARIS

Résumé. Dans le modèle de diffusion considéré, la fonction de dérive est lipschitzienne sur $\mathbb{R}$ mais non dérivable en r_0. Notre but est d'estimer ce paramètre de seuil à partir d'une observation discrétisée à pas h_n du processus stationnaire et ergodique. Les moindres carrés associés au schéma d'Euler fournissent un estimateur consistant de r_0 lorsque $nh_n \to +\infty$ et $h_n \to 0$. Si de plus, $nh_n^3 \to 0$, l'estimation est asymptotiquement normale avec la vitesse standard $\sqrt{nh_n}$.

Abstract. We consider a diffusion model for which the drift b is Lipschitz continuous on $\mathbb{R}$ but its derivative is not continuous in r_0. Our aim is to estimate this threshold parameter from discrete observations at step h_n of the stationary and ergodic process. The least squares method based on the approximate discrete-time Euler's scheme provides a consistent estimator when $nh_n \to +\infty$ and $h_n \to 0$. Moreover if $nh_n^3 \to 0$, this estimator is asymptotically normal with a standard rate of order $\sqrt{nh_n}$.

1 Introduction

Soit le modèle de diffusion défini sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbf{P})$:

$$dX_t = b\,(X_t - r_0)\,dt + dB_t, \quad X_0 = \xi \qquad (E_{r_0})$$

La fonction b est continue sur $\mathbb{R}$ et de classe C^2 sur $\mathbb{R} - \{0\}$ telle que, si l'on note b' sa dérivée première, $\lim_{x \to 0^{\pm}} b'(x) = b'\left(0^{\pm}\right)$ existent et sont finies avec $b'(0^+) \neq b'(0^-)$. Nous dirons alors que b est un élément de $C^2_{\{0\}}$ et nommerons **diffusion à seuil** un tel modèle. Sous l'hypothèse $b \in C^2_{\{0\}}$, le modèle étudié appartient à la classe plus générale des modèles réguliers : $x \to b\,(x, r)$ continue et différentiable en moyenne quadratique par rapport à r. Dans [7], Kutoyants montre que, pour un modèle régulier, l'estimateur du maximum de vraisemblance (E.M.V.) associé à une observation continue du processus stationnaire ergodique sur $[0, T]$ est consistant, asymptotiquement normal à la vitesse standard $\sqrt{T}$ et asymptotiquement efficace au sens des estimateurs asymptotiquement minimax selon le critère de Hajek [4] $(T \to +\infty)$. Notre but est de fournir, dans le cas d'une diffusion à seuil, une estimation de r_0, consistante et asymptotiquement normale avec une vitesse et une variance asymptotique identiques à celles de l'estimateur du maximum de vraisemblance, mais basée sur une observation discrétisée à pas fin h_n $(h_n \to 0)$ du processus, $(X_{kh_n})_{k=0,n}$, avec $nh_n = T$. Nous nous basons pour cela sur l'estimateur des Moindres Carrés (M.C.) associés au schéma d'approximation d'Euler.
Lorsque $(X_t)_{t \geq 0}$ est ergodique et la fonction de dérive est de classe C^2 sur $\mathbb{R}$, les résultats de Florens-Zmirou [3] entrainent que cet estimateur est consistant, lorsque $nh_n \to \infty$ et $h_n \to 0$. Si de plus $nh_n^3 \to 0$, il est asymptotiquement normal à la vitesse $\sqrt{nh_n}$ et sa variance asymptotique est égale à l'inverse de l'information de Fisher. Nous montrons que les propriétés asymptotiques de cet estimateur sont préservées lorsque $b \in C^2_{\{0\}}$ si l'on suppose que b est de Lipschitz sur $\mathbb{R}$. Nous établirons également que cet estimateur est asymptotiquement équivalent à l'E.M.V. basé sur une observation continue de la trajectoire sur $[0, T]$.

2 Propriétés asymptotiques

On définit $L_0^2(\mu) = \left\{f : \mathbb{R} \to \mathbb{R},\ \mu(f) = 0,\ \mu\left(f^2\right) < \infty\right\}$ et pour tout $p \in \mathbb{N}^*$, $\|f\|_{p,\mu} = \left[\mu\left(|f|^p\right)\right]^{\frac{1}{p}}$.
On suppose que $r_0 \in]a,b[$, où a et b sont deux réels tels que $a < b$. Posons les conditions suivantes **(A)** :

(A 1) b appartient à $C^2_{\{0\}}$ et il existe $K > 0$ telle que : $\forall (x,y) \in \mathbb{R}^2, |b(x) - b(y)| \leq K|x-y|$.
On suppose de plus que b'' est à croissance polynomiale.
Sous (A 1), pour tout $r \in \mathbb{R}$, (E_r) admet une unique solution forte $X^r = (X_t^r)_{t\in[0,T]}$. Notons P_r^T la loi de ce processus.

(A 2) Notons $s_{r_0}(x) = \exp\left(-2\int_{r_0}^x b(y-r_0)\,dy\right)$ la dérivée de la fonction d'échelle associée à (E_{r_0}) et supposons que : $\int_0^\infty s_{r_0}(x)\,dx = \int_{-\infty}^0 s_{r_0}(x)\,dx = \infty$ et $\int_{-\infty}^\infty \left[s_{r_0}(x)\right]^{-1} dx = C(r_0) < \infty$.
Notons μ_{r_0} la loi de densité $q_{r_0}(x) = \left[C(r_0)\,s_{r_0}(x)\right]^{-1}$ par rapport à la mesure de Lebesgue.

(A 3) La loi de ξ est μ_{r_0}.
Sous (A 1), (A 2) et (A 3), $X^{r_0} = (X_t^{r_0})_{t\in[0,T]}$ est stationnaire stricte, de loi invariante μ_{r_0}, et ergodique.

(A 4) Notons $\pi_{r_0}^t$ le semi-groupe associé à (E_{r_0}). On suppose que μ_{r_0} admet des moments de tous ordres et qu'il existe $\delta > 0$ tel que pour tout $t \geq 0$, on a : $\forall f \in L_0^2(\mu_{r_0})$, $\left\|\pi_{r_0}^t f\right\|_{2,\mu_{r_0}} \leq \exp(-\delta t)\,\|f\|_{2,\mu_{r_0}}$

Sous (A), $r \to b(\cdot - r)$ est dérivable au point r_0 dans $L^2(\mu_{r_0})$, de dérivée $-\dot{b}(\cdot - r_0)$ où $\dot{b}(y) = b'(y^+)\,I_{(y>0)} + b'(y^-)\,I_{(y\leq 0)}$. L'information de Fisher asymptotique au point r_0 associée à la famille de lois $\left(P_r^T\right)_{r\in]a,b[}$ est ainsi donnée par : $I(r_0) = \int_{\mathbb{R}} \left[\dot{b}(x-r_0)\right]^2 q_{r_0}(x)\,dx$. Sous (A), on vérifie que : $I(r_0) > 0$.

La fonction d'estimation couplant la méthode des M.C. au schéma d'Euler est donnée par :

$$U_n(r) = \frac{1}{nh_n^2}\sum_{k=1}^n \left(X_{kh_n} - X_{(k-1)h_n} - h_n b\left(X_{(k-1)h_n} - r\right)\right)^2$$

Sous (A), $r \to U_n(r)$ est le processus de contraste relatif à la fonction de contraste $r \to U(r,r_0) = \int_{\mathbb{R}} (b(x-r) - b(x-r_0))^2\, q_{r_0}(x)\,dx$ (définition 3.2.7, [2]). On note alors $\widehat{r}_n$ l'estimateur du minimum de contraste (E.M.C.) défini par : $\widehat{r}_n = \inf_{r\in[a,b]} U_n(r)$

2.1 Consistance

Dans la suite, on note : $\forall t \in [0,T]$, $X_t^{r_0} = X_t$. Soit $\mathcal{F}_t$ la tribu engendrée par $(X_s)_{0\leq s\leq t}$ et $\mathbf{E}$ l'espérance sous $\mathbf{P}$.
On définit, pour tout $n \in \mathbb{N}^*$, la martingale centrée M_n et le reste R_n :

$$M_n(r) = \sum_{k=1}^n b\left(X_{(k-1)h_n} - r\right)\left(X_{kh_n} - \mathbf{E}\left[X_{kh_n} \mid \mathcal{F}_{(k-1)h_n}\right]\right)$$

$$R_n(r) = \sum_{k=1}^n \left[b\left(X_{(k-1)h_n} - r_0\right) - b\left(X_{(k-1)h_n} - r\right)\right]\left[\mathbf{E}\left[X_{kh_n} \mid \mathcal{F}_{(k-1)h_n}\right] - X_{(k-1)h_n} - h_n b\left(X_{(k-1)h_n} - r_0\right)\right]$$

L'accroissement $U_n(u) - U_n(v)$ s'exprime en fonction de ceux de M_n et R_n. Utilisant la propriété de Lipschitz de b, on a en particulier : $\sup_{|v-u|\leq\eta,(v,u)\in[a,b]^2} |U_n(u) - U_n(v)| \leq V_n(\eta)$, où

$$V_n(\eta) = \frac{2}{nh_n}\sup_{|v-u|\leq\eta,(v,u)\in[a,b]^2}|M_n(v) - M_n(u)| + \frac{2}{nh_n}\sup_{|v-u|\leq\eta,(v,u)\in[a,b]^2}|R_n(v) - R_n(u)| + 2K^2(b-a)\eta$$

Or, sous (A), on montre que : $\forall \eta \in [0, b-a]$, $\mathbf{E}\left[\sup_{|v-u|\leq\eta,(v,u)\in[a,b]^2} |R_n(v) - R_n(u)|\right] \leq C\ \eta\ nh_n^2 K$.
De plus, si $(h_n)_{n\geq 0}$ est bornée, il existe C telle que pour tout couple $(u,v) \in [a,b]^2$ et tout $n \in \mathbb{N}^*$:
$\mathbf{E}\left[(M_n(u))^2\right] \leq Cnh_n$, $\mathbf{E}\left[(M_n(v) - M_n(u))^2\right] \leq Cnh_n(v-u)^2$. Par une adaptation du théorème 19 de Ibragimov et Has'minskii ([5], annexe I, p.372), on obtient alors : $\forall \eta \in [0, b-a]$,
$\mathbf{E}\left[\sup_{|v-u|\leq\eta,(v,u)\in[a,b]^2} |M_n(v) - M_n(u)|\right] \leq \widetilde{C}\sqrt{\eta}\sqrt{nh_n}$ où $\widetilde{C}$ est une constante finie indépendante de n.
Donc : $V_n(\eta) \xrightarrow{L^1(\mathbf{P})} 2K^2(b-a)\eta$
La consistance de $\widehat{r}_n$ découle alors directement de l'application du théorème 3.2.8 de [2] :

Théorème 1 *Sous (A), si $nh_n \to +\infty$ et $h_n \to 0$ alors $\widehat{r}_n \xrightarrow{\mathbf{P}} r_0$.*

2.2 Vitesse de convergence et Normalité asymptotique

Puisque la fonction de dérive n'est pas dérivable en r_0, ces résultats ne peuvent pas être obtenus par la méthode classique basée sur le développement de Taylor de la dérivée du contraste. Ainsi, après avoir identifié la vitesse de convergence de $\widehat{r}_n$ vers r_0, nous développons directement le contraste autour de r_0 en mettant en évidence une décomposition qui satisfait une propriété LAN (Local Asymptotic Normality, cf. [5]).

L'identification de la vitesse de convergence de $\widehat{r}_n$ vers r_0 repose sur la transposition d'arguments développés par Chan [1] dans le cadre de modèles autorégressifs à seuil à temps discret.

Théorème 2 *Sous (A), si $(h_n)_n$ est telle que $nh_n \to \infty$ et $\left(nh_n^3\right)_n$ est bornée, alors pour tout $\varepsilon > 0$, il existe $H_0 > 0$ et $n_{H_0} \in \mathbb{N}^*$ tels que pour $n \geq n_{H_0}$:* $\mathbf{P}\left(|\widehat{r}_n - r_0| \leq \frac{H_0}{\sqrt{nh_n}}\right) \geq 1 - \varepsilon$

Si $nh_n \to +\infty$ et $\left(nh_n^3\right)_n$ est bornée, la suite $\left(\sqrt{nh_n}\,(\widehat{r}_n - r_0)\right)$ est tendue. On introduit alors : $Z_n(u) = nh_n\left(U_n\left(r_0 + \frac{u}{\sqrt{nh_n}}\right) - U_n(r_0)\right)$. Généralisant la formule de Taylor dans le cas d'une fonction de $C^2_{\{0\}}$, nous exhibons une décomposition de Z_n de la forme : $Z_n(u) = 2I(r_0)\,u\Delta_n + I(r_0)\,u^2 + \Psi_n(u)$, où $I(r_0)$ est l'information de Fisher définie précédemment et $\Delta_n = \frac{I(r_0)^{-1}}{\sqrt{nh_n}} \sum_{k=1}^n \dot{b}\left(X_{(k-1)h_n} - r_0\right)\left[X_{kh_n} - \mathbf{E}\left[X_{kh_n} \mid \mathcal{F}_{(k-1)h_n}\right]\right]$.

Nous appuyant sur le théorème de limite centrale pour les suites triangulaires (théorème 2.8.43., [2]), on montre que Δ_n converge en loi vers la loi Normale $\mathcal{N}\left(0, I(r_0)^{-1}\right)$. Par ailleurs, sous (A), Ψ_n est une fonction presque sûrement continue telle que, si $nh_n \to \infty$ et $nh_n^3 \to 0$, on a, pour tout $B > 0$, $\sup_{|u|\leq B} |\Psi_n(u)| \xrightarrow{\mathbf{P}} 0$. Le reste $\Psi_n(u)$ est la somme de cinq termes dont $R_n\left(\frac{u}{\sqrt{nh_n}}\right)$. Le résultat de convergence est obtenu en majorant le $\sup_{|u|\leq B}$ de la somme par la somme des cinq $\sup_{|u|\leq B}$ et en montrant que chacun de ces termes tend vers 0 au sens de la convergence $L^1(\mathbf{P})$. Pour le terme $R_n\left(\frac{u}{\sqrt{nh_n}}\right)$, on utilise la majoration suivante : $\mathbf{E}\left[\sup_{|u|\leq B}\left|R_n\left(\frac{u}{\sqrt{nh_n}}\right)\right|\right] \leq O\left(\sqrt{nh_n^3}\right)$. Le résultat de convergence est ainsi obtenu par le renforcement de l'hypothèse de décroissance de $(h_n)_n$ vers 0.

La propriété satisfaite par la fonction Ψ_n est plus forte que celle requise pour établir le caractère LAN du contraste au point r_0. On adapte alors le théorème 1.2 de Ibragimov et Has'minskii [5] : ce résultat assure la normalité asymptotique de l'estimateur du maximum de vraisemblance lorsque la vraisemblance du modèle considéré satisfait, en outre, une condition LAN. On obtient ainsi le résultat de convergence en loi de $\widehat{r}_n$:

Théorème 3 *Sous (A), si $(h_n)_n$ est telle que $nh_n \to \infty$ et $nh_n^3 \to 0$, on a :*

$$\sqrt{nh_n}\,(\widehat{r}_n - r_0) + \Delta_n \xrightarrow{\mathbf{P}} 0, \qquad \sqrt{nh_n}\,(\widehat{r}_n - r_0) \xrightarrow{\mathcal{L}(\mathbf{P})} 0$$

3 Comparaison de l'E.M.V. et de l'E.M.C.

Un E.M.V., $\widehat{r}_T$, basé sur une observation continue sur $[0,T]$ du processus solution de (E_{r_0}) est défini par :

$$-L_T(\widehat{r}_T) = \inf_{r\in[a,b]}(-L_T(r)), \; L_T(r) = \int_0^T [b(X_t - r) - b(X_t - r_0)]\, dX_t - \frac{1}{2}\int_0^T \left[b^2(X_t - r) - b^2(X_t - r_0)\right] dt$$

Sous (A), $\left(P_r^T\right)_{r\in]a,b[}$ est LAN au point r_0 et le terme principal du développement local de $L_T(r_0) - L_T\left(r_0 + \frac{u}{\sqrt{T}}\right)$ est $\Delta_T = \frac{I(r_0)^{-1}}{\sqrt{T}} \int_0^T \dot{b}(X_t - r_0)\,[dX_t - b(X_t - r_0)\,dt]$ qui converge vers une loi $\mathcal{N}\left(0, I(r_0)^{-1}\right)$ (théorème 3.3.8,[7]).

Nous montrons que $\widehat{r}_T$ et $\widehat{r}_n$ sont asymptotiquement équivalents en nous basant sur les trois résultats de convergence suivants :

1. lorsque $T \to +\infty$, $\sqrt{T}(\widehat{r}_T - r_0) + \Delta_T \xrightarrow{\mathbf{P}} 0$. Ce résultat est une conséquence du théorème 3.4.4. de [7].
2. lorsque $T = nh_n \to +\infty$ et $nh_n^3 \to 0$, $\sqrt{T}(\widehat{r}_n - r_0) + \Delta_n \xrightarrow{\mathbf{P}} 0$ (théorème 3).
3. lorsque $T = nh_n \to +\infty$ et $h_n \to 0$, $\Delta_n - \Delta_T \xrightarrow{\mathbf{P}} 0$. Ce résultat est obtenu en montrant que $\mathbf{E}\left[\left(\dot{b}(X_h - r_0) - \dot{b}(X_0 - r_0)\right)^2\right] = o(1)$.

Il suffit alors d'écrire : $\sqrt{T}(\widehat{r}_T - \widehat{r}_n) = \left(\sqrt{T}(\widehat{r}_T - r_0) + \Delta_T\right) + (\Delta_n - \Delta_T) - \left(\sqrt{nh_n}(\widehat{r}_n - r_0) + \Delta_n\right)$. De ce qui précède, on déduit :

Théorème 4 *Sous (A), lorsque $T = nh_n \to +\infty$ et $nh_n^3 \to 0$,* $\sqrt{T}(\widehat{r}_T - \widehat{r}_n) \xrightarrow{\mathbf{P}} 0$

References

[1] Chan, K. S. (1993) *Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model.*- The Annals of Statistics, Vol. 21, No. 1, 520-533.

[2] Dacunha-Castelle, D. & Duflo, M. (1993) *Probabilités et Statistiques : Problèmes à temps mobile.*- 2^{eme} Ed., Masson.

[3] Florens-Zmirou, D. (1989) *Approximate discrete time schemes for statistics of diffusion processes.*- Statistics 20, 547-557.

[4] Hajek, J. (1972) *Local asymptotic minimax and admissibility in estimation.*- Proc. 6^{th} Berkeley Symp. on Math. Statist. and Prob. 1, Univ. of Calif. Press, 175-194.

[5] Ibragimov, I. A. & Has'minskii, R. Z. (1981) *Statistical Estimation- Asymptotic Theory.*- Springer-Verlag.

[6] Kessler, M. (1997) *Estimation of an ergodic diffusion from discrete observations.*- Scand. J. Stat. 24 (2), 211-229.

[7] Kutoyants, A. Y. (1984) *Parameter estimation for stochastic processes.*- Research and Exposition in math. 6, Heldermann Verlag, Berlin.